~ƀ~

Carbone based living forms have been evolved to contain a set of proteins with variable physical>~<chemical properties. Self-replicating organic living organisms can be referenced to a {genome} that is able to produce a [set] of proteins. The diversity of living entities, from unicellular organisms to multicellular mammals can be corresponded to the varying repertoire of proteins.

The [ribosome] is a macromolecular complex which is responsible for protein synthesis in all known living cells according to corresponded genetic information. Ribosomes are organelles ubiquitously produced in all cells and tissues able to catalyze a protein synthesis.

Ribosomal proteins (RPs) genes encode structure components of ribosomes. Ribosomes carry out one of the most fundamental processes in living systems translating genetic information from mRNA into proteins. In eukaryotes, each [ribosome] consists of two subunits, the <small> subunit and <large> subunit, consisting of different RPs and different types of RNAs.

Types of [RNA]:

1. **mRNA**- Messenger RNA: encodes amino acid sequence of a polypeptide; a single stranded molecule of RNA that can be read by [ribosome] in the process of protein synthesis | created by RNA polymerase, can be corresponded to the genetic sequence of a gene;
2. **tRNA**- Transfer RNA: function at specific sits in the [ribosome] during protein synthesis from an mRNA molecule. The [ribosome] express three binding sites for tRNA_<attached amino acid> that are introduced to [ribosome] by ~<elongating factors> | vary in number of tRNA genes organelles. In observed eukaryotes there are 22 mitochondrial tRNA genes, transcribed by <RNA Polymerase III>. Polymerase recognizes in tRNA molecule two highly conserved promoter sequences, first promoter begins at [+8] of mature tRNAs, the second promoter is located [30-60] nucleotides downstream of the first promoter.
3. **rRNA**- Ribosomal RNA: the environmental, physical, mechanical, semantic factor of tRNA~mRNA translation into proteins; makes up the ribosomes, the organelles that translate the mRNA:
4. **snRNA**- small nuclear RNA: with proteins forms complexes that are involved in RNA processing in eukaryotes (not observed in prokaryotes);
5. **mtRNA**- Mitochondrial RNA: genome of bacterial origin, primary encoding set of vital subunits of the energy generating systems: oxidative phosphorylation and electron transport chains; pathways that generate Adenosine Triphosphate ~[ATP]~ the primary energy substrate of a eukaryotic cells;
6. . .

A [proteome set] guide the morphology and functionality of an organism.

RP genes are highly conserved in all domains of life. Each peptide was found to have unique amino acid sequence corresponded with very limited to none similarities between each other.

Genes encoding ribosomal proteins are expressed at rate limiting levels, rendering their biological functions as highly sensitive to a/the copy-number variations that results in genomic instabilities.

/# cells with a reduced number of ribosomal protein genes are eliminated in cell-to-cell competition with wild type cells #/

<Rib~0~Proteins> are able 2 subject [ribosome] to feedback regulation;

There are several examples of post-transcriptional autoregulation by yeast ribosomal proteins, often through inhibition of splicing necessary for protein expression.

The one of fundamental characteristics in feedback regulation of gene expression is that the activity of a gene product, which is usually proportional to its concentration, is higher than the set point, or “desired” level for cell, then its expression is altered.

Numerous studies suggest that various genes in different kinds of organisms are transcribed discontinuously, i.e., in short burst or pulses with periods of inactivity between them. Data indicates that the/a ribosomal genes in the human cells are transcribe discontinuously with periods ranging with dynamics of rDNA transcription that follows the pattern, in which, the bursts of activity are alternated by periods of transcription events.

Gene expression can be seen as a multistage process in which information can be represented in form of DNA and used to generate gene products. Gene expression can be seen as a process that involve at least two distinguishable stages: transcription and translation;

- During transcription: RNA copies of DNA genes are synthesized by enzymes <RNA Polymerase>~s. The product is the messenger RNA (mRNA), which can be seen as series of nucleotide triplets (codons), which codes the order of amino-acid needed to be combined to synthesize the protein;

- During translation: the information in the mRNA is decoded and the protein is synthesized. Complex macromolecules – [ribosomes] – introduced to the start codon in the mRNA, sequentially decode each codon to its corresponding amino-acid. The amino-acid peptide is elongated until the [ribosome] reaches a stop codon, detaches from the mRNA and the resulting amino-acid-chain/peptide is released and becomes a functional protein;

The transcription-translation complex has a dual role: to synthesize copies of itself as well as peptides involved in the post translation protein synthesis. Two key teem players in this process, {<RNA Polymerase>|s and [ribosome]|s} are able to self-replicate by jointly producing sub-components which subsequently self-assemble to new RNA Polymerases and ribosomes.

Under the rich medium growth condition, the majority of <RNA polymerase>~s clusters actively engage in ribosomal RNA transcription. These clusters presence and spatial distribution did not appear to require a high level of rRNA transcription activity or the presence of multiple operons. Data suggest that existence of clusters were likely influenced by the underlying nucleotide structure.

Ribosome assembly consists of both sequential and parallel processing steps. The mixture of sequential and parallel elements provides a rich environment for potential assembly pathways. Under a particular growth condition, [ribosome] assembly flux can partition between the available parallel assembly pathways according to the rates of those pathways, with greater flux along the more rapid pathways. The availability of alternative parallel pathways provides an added degree of robustness to buffer the system against changes in proteins or assembly factors availability.

[Ribosome] biogenesis is a process involving three RNA polymerases;

[<RNA Polymerase I>~<RNA Polymerase II>~<RNA Polymerase III>~ addition of/to post/non-ribosomal factors required for the protein synthesis, processing, transportation and assembly of pre-ribosomes];

general [ribosome] biogenesis can be described as multistep process:

1. 5S rRNA generated by <RNA Polymerase III>,
2. 47S/45S rRNA synthesized by <RNA Polymerase I>,
3. RP-encoding mRNA produced by <RNA Polymerase II>,
4. along with non-ribosomal factors and small nucleolar RNAs,

Steps [1-4] are assembled to yield the 90S pre-ribosomes in the nucleus, which undergo multiple modifications and subsequent separation into pre-60S and pre-40S particles. During transport from the nucleus to the cytoplasm, these pre-ribosomes are dissociated from most of their non-ribosomal factors and maturated to <60S> and <40S> subunits for protein translation.

Ribosome biogenesis is rapid requiring about 2 minutes for production of a single [ribosome]; the vast majority of assembly events resulting in mature translationally active complexes.

In bacteria, a ribosome is a particle size about 250A° in diameter consisting mainly of two subparts: <30S> and <50S> subunits, together forming the <70S> subunit.

/# A°-angstrom = 0.1 nm, the unit “S” stands for Svedberg, which is a measure of the sedimentations rate #/

The <30S> subunit plays a role in decoding mRNA by monitoring base pairing between codon and anticodon. The <30S> (average number of atoms 51742 | mass fractal dimension [2.82 ⌀ 0.07(st.dev)]) subunit could be further subdivided into:

- <16S> rRNA molecule (comprised of approximately 1500 nucleotides, average number of atoms 32514 | mass fractal dimension [2.58 ⌀ 0.06(st.dev)]);
- Addition ~ about 20 different proteins;

The <50S> subunit catalyzes peptide bond formation between the incoming amino acid and the nascent peptide chain. The <50S> subunit (average number of atoms 62673 | mass fractal dimension [3.07 ⌀ 0.08(st.dev)]) could be subdivided into:

- <5S> rRNA subunit;
- <23S> rRNA subunit (comprised of approximately 3000 nucleotides, average number of atoms 59017 | mass fractal dimension [3.11 ⌀ 0.07]);
- Addition ~ about 30 different proteins;

Each of a subunit composed of one of four different basses (denoted as A, C, G and U) with sugar phosphate backbones.

Accessory biogenesis factors help guide the [ribosome] assembly process, many of them are nonessential, indicating that the reaction they catalyze can be bypasses in an alternative pathway. The factors are present and could be necessary under particular environmental conditions. Ribosome profiling experiments as well as direct quantitative mass spectrometry measures suggest that the assembly factors constitute up to 5% of total ribosomal mass.

Experiments demonstrates that Archaeal NLStype motifs can substitute eukaryotic NLSs to direct intracellular transport of ribosomal proteins in eukaryotic cells.

A typical nuclear localization signal is a short and surface exposed stretch of basic residues that is recognized by specialized receptors, which transfer NLS-containing proteins across the nuclear pores;

Available data indicates that <RNA polymerase>|~s are spatially localized into distinct clusters in *E.Coli* and *B.Subtilis* cells. Existence of spatially organization for molecular components in prokaryotic systems imply possibility in compartmentalization without use of membranes.

Bacterial cells possess a remarkable degree of spatial organization of cellular components with activities without use of a membranes. In *E.Coli* and *B.Subtilis* cells growth, <RNA polymerase>|~s was found to form dense foci simultaneously instead of distributing homogenously within a cell.

Fossil records suggest that a group of ancestral prokaryotic cells were transformed into what now referred to as eukaryotes, they acquired a DNA-storage compartment – the nucleus – that was separated from cytoplasm by a nuclear membrane equipped with selectively penetrable channels, the nuclear pores. To assist in passage via the nuclear pores between cytoplasm and the nucleus, cellular proteins have evolved specialized signal sequences, the nuclear localization signals (NLSs).

Investigation to the origins of nuclear localization signals in eukaryotic calls revealed that protein segments that were describe as NLSs in eukaryotic ribosomal proteins were also present in homologues proteins from Archaea`s. The finding reveals a group of rare evolutionary intermediates – NLStype motifs in archaeal ribosomal proteins. This finding indicates that at least some of NLSs were present in proteins considerably earlier than the event when cells separated into the nucleus and cytoplasm or presenting an evidence of causality invariance for evolution in action.

/# in permissive environments, *E.Coli* can double dry mass every 21 minutes. During this time ribosomes, RNA Polymerases and proteome are all doubled; #/

/# in cell, multiple translation processes take place concurrently, utilizing shared translation resources. *An* estimate: in a yeast, cell contains about 60000 mRNA molecule and about 240000 ribosomes #/

/# The mass fractal dimension is defined as the number of monomers (atoms in this case), N, enclosed in a sphere of radius R. It can be used as a measure of the compactness as follows: relation between the average radius and the mass of the protein chains can be described by [mass**X**fractal] dimension>~<

The mass fractal dimension exhibits the molecule`s space-filling ability, the larger is Dm, the more units are in the sphere. When fractal dimension is less than 3, the structure can be seen as having “empty” or “void” space. The mass fractal dimension of the protein was shown to lie near 3, suggesting a compact three-dimensional object.

Studies suggest that the fractal dimension of less than 3 may be intrinsic and universal characteristics of the protein chain. #/

Multiple [ribosome]|s may decode the same mRNA molecule simultaneously;

The ribosomal proteins express the autoregulation, proportional to concentration, as well as self-assembly flux.

The rapid glucose consumption was facilitated by the increased glycolysis flux and more efficient glucose transport across cellular membranes.

Evolutional topology tree suggest that the RP genes have been independently duplicated in each species after their divergence from common ancestor.

Among publicly available data on 285 fungi, ten species express massive RP duplication with more than 50% of RP families (>40) have duplicated copies.

Available data suggest that the retention genes might have been driven by their contribution to the evolution of strong fermentative ability in these organisms.

Duplication of RP genes are able to provide an additional substrate on which <RNA Polymerase II> can transcribe into <RibOProteins>, increasing selected [ribosome] biogenesis.

Most eukaryotic species fully oxidize glucose, their primary carbon and energy source, through mitochondrial oxidative phosphorylation in presence of oxygen for maximum energy production. In contrast some budding yeasts and fission yeasts predominantly ferment sugar to ethanol in presence of excess sugars, even under aerobic conditions. /# aerobic fermentation #/

Fermentation is less efficient way to generate energy. Through fermentation, each glucose molecule yields 2 ~[ATP]~ from glycolysis, compared to 32 ~[ATP]~ through mitochondrial oxidative phosphorylation pathway. In sugar rich environments, fermentative organisms are able to produce more ~[ATP]~ per unit time by rapidly consuming sugars through fermentation, providing selective advantages to these organisms.

Gene duplication has served as a driving force for the evolution of new phenotypic traits and contributed to adaptation of organisms to their specific niches. The available data suggests that the increase RP expression by gene duplication might have enhanced these organism’s ability to rapidly consume glucose through the fermentation pathway, providing competitive advantages to them.

Across all eukaryotic species, RP genes are the group of most abundantly transcribes genes in eukaryotic cells, accounting for 50% of <RNA Polymerase II> transcription. RP genes have the highest density of bound <RNA Polymerase II>: 100RP genes have on average more than 60% of the maximum <RNA Polymerase II> occupancy, while a majority of the genome has less than 5% of the maximum <RNA Polymerase II> density.

/# To understand the role of assembly factors in [ribosome] biogenesis, a series studies have used genetic and/or pharmacological assembly factor inhibitors. Cells are grown under factor-restricted conditions, which results in the accumulation of apparent assembly intermediates whose composition and structure is then determined. Using this approach, researchers have suggested roles for assembly factors in aiding in docking of entire structural domains, properly positioning of proteins on a [ribosome], or in blocking association until maturation is complete. #/

It has been demonstrated that increased copy number of genes related to glycolysis and glucose transporters have played an important role in switch of glucose metabolism. It is reasonable to propose that increased RP gene dosage increases abilities of organisms in specific [ribosome] biogenesis, resulting in increased biogenesis of glycolysis enzymes, glucose transporters, and other building blocks for cell growth and proliferation. Massive duplication of RP genes made it possible to more rapidly consume glucose through the low-efficient fermentative pathway, and at the same time maintain a high growth rate. Such physiological chrematistics could possibly provide selective advantages to the organisms in sugar rich environments.

The ribosomal proteins are RNA binding proteins, having as their primary target the rRNA of the [ribosome]. To become autoregulatory, the mRNA sequence mimic of the primary target site sequence with lower affinity, as consequence binding of the rRNA is saturated before the regulatory site become bound by the peptide.

Feedback regulation of protein expression requires that when the activity of the protein is higher than the homeostasis point of the cell, its expression is reduced, and when the activity is lower than that point its expression is increased. This process allows the cell to maintain expression in a narrow range around its set point.

Efficient [ribosome] biogenesis consumes bigger part of cellular energy in form of ATP and thus tightly coupled with the energy status of the cell. This feature makes the nucleolar process highly sensitive to nutrient deprivation, as it was demonstrated by several studies on the target of rapamycin <TOR> signaling that has been shown to play a key role in linking the cellular nutrient status to [ribosome] biogenesis.

In yeast, the transcription of RP genes is regulated by <TOR>mediated regulation of Forkhead transcription factor <FHL1> along with its co-activator <IFH1> and the repressor <CRF1>.

- Under favorable conditions, <FHL1> binds to RP gene promoter regions and activates their transcription;
- Under unfavorable conditions, the <FHL> bind to its repressor <CRF1>, which leads to the suppression of RP gene transcription;

The target of Rapamycin (TOR) protein kinase reprograms cellular metabolism under various environmental stresses. The overexpression of <TOR> in *Arabidopsis* resulted in increased plant growth including yield and biomass when compared with the wild type under both controlled and limited water conditions.

Target of Rapamycin <TOR> is a conserved eukaryotic Serine/Threonine protein kinase, which controls signaling networks involved in cell growth and development; is a key regulator of cellular metabolism in plants and mammals; modulates growth and development in association with the nutrient availability and energy status of the cell as well as plays a central role in the regulation of low energy signaling and metabolic reprograming under stress conditions.

The <TOR> protein exhibits five conserved domains:

1. <HEAT> repeat domain; /# involved in rRNA synthesis #/
2. <FAT> domain; /# involved in scaffolding and protein-protein interactions #/
3. <FRB> domain; /# provides the building site for a macrocyclic, immunosuppressant drug - <Rapamycin> #/
4. Serine/Threonine kinase domain;
5. <FATC> domain; /# involved in scaffolding and protein-protein interactions #/

<TOR> protein and their interacting partners function as two complexes [TORC1] and [TORC2]:

> [TORC1] is a complex of [<TOR>~<LST8>~<RAPTOR>] that regulates basic functions for cell survival such as cell growth, ribosome biogenesis, protein translation in response to nutrients and energy;
>
> [TORC2] is a complex of [<TOR>~<LST8>~<SIN1>~<RICTOR>] that regulates cytoskeletal structure, <actin> polarization, cell polarity as well as others functions;

Studies in yeast showed that inhibition of <TOR> globally repress the transcription of RP-encoding genes.

<TOR> signaling pathway promotes [ribosome] biogenesis by regulating a/the production of ribosomal proteins and rRNA through multiple mechanisms;

<TOR> was found to booster rDNA transcription by sequestering/catching <TIF-IA> in the nucleus and facilitating the formation of the <RNA Polymerase I> initiation complex.

Rapamycin has also been shown to reduce the activity of <RNA Polymerase III> and its specific transcription factor <TF-IIIB>, repressing the transcription of the 5S rRNA genes.

/# Downstream target of the <TOR>pathway>~<AMP_activated protein kinase> (AMPK); can be activated upon glucose deprivation to phosphorylate and inactivate <TIF-IA>, leading to the repression of <RNA Polymerase I> mediated transcription #/

/# Plant ribosomal proteins present large class of conserved and abundant RNA binding proteins with functions in stabilizing rRNA structures providing the efficient tool for protein synthesis. They are often encoded by small multigene families comprising up to eight active genes, suggesting potential specialization or extra-ribosomal activities for some family members. *Arabidopsis* plants with overexpressed <TOR> exhibiting enhanced tolerance to the osmotic and salt stress treatments #/

Examples: Two copies of <RPL11> genes are found in each *Schizosaccharomyces* species;

Similar patterns are present for the <RPS19> gene, which encodes a ribosomal small subunit protein.

<RPP1>, <RPP2>, <RPL9> and <RPS22> families have duplicates in most budding yeasts, suggesting that they have been duplicated before divergence of all budding yeasts.

Two copies of <RPL6> gene are present in budding yeasts and in all four fission yeast species. More copies of <RPL6> gene are presented in some [*Mucoromycota*] species, [*Phycomyces blakesleeanus*], [*Lobosporangium transversal*], [*Rhizopus microspores*], and [*Rhizopus delemar*], which have 2,3,4 and 5 copies of <RPL6> gene respectively.

Survey of 295 fugal genomes revealed that massive duplications of RP genes are not prevent. A significant increase in RP gene copy numbers had independently occurred in small number of species in three distantly related lineages of fungi. The data suggest that RP genes in the fission yeasts were likely individually duplicated by small-scale duplication events.

During interlocus gene conversion, one gene serves as a DNA donor that replaces the sequence of paralogue gene. As a result, the sequence of paralogues genes has been homogenized /# bitter milked | can be achieved by spinning or by passing through some needle neck #/, resulted in duplication event sometimes referred as “concerted evolution”.

If genomic DNA is the donor, only the exon sequence of acceptor is replaced. In most observed cases, only the coding sequence have been homogenized by gene conversion, supporting cDNA as the probable gene conversion donor.

During gene conversion, the genomic sequence of the ‘acceptor’ locus is replaced by a ‘donor’ sequence through process of recombination.

The donor can be genomic DNA or complementary DNA (cDNA) derived from an mRNA intermediate.

/# Concerted evolution is a genetic process by which repetitive DNA sequences are homogenized so than the sequence of a given multigene family within a species share higher sequence identity of the homologue’s genes compared to a family of different species. Events of concerted evolution observed in all living organisms, from bacteria to humans. The pattern is – repetitive sequences express higher level of similarity among members of a family withing the species than between a species #/

Mechanism example:

Translation is initialized on the 3~AUG~5 start codon on the mRNA. In eukaryotes, this codon is recognized by a pre-initiation complex of the small subunit, initiator tRNA and several initiation factors <IF>s.

Translation initiation factor <eIF5> replaces <eIF1> and the ribosomal <40S> subunit to promote start-codon recognition.

/# <eIF5> is a GTPase activating protein of <eIF2> that additionally promotes stringent 3~AUG~5 selection #/

The affinity of <eIF1> for <40S> subunit is > 30-fold higher than that of the <eIF5>, which favor initial binding of <eIF1> over <eIF5>.

The affinity of <eIF1> for the <43S>mRNA(AUG) complex is 20-fold lower, whereas the affinity of <eIF5> for the same complex is more than >300-fold higher, in comparison to their respective affinities for the <40S> subunit.

As a result, following 3~AUG~5 recognition, the affinity of <eIF5> for the preinitiation complex exceeds that of <eIF1> by two orders of magnitude.

Underling mechanism is most likely can be general as one of the most fundamental in nature;

+ can be seen~used as a form for/of intelligent consensus sequence design

ь

1. https://arxiv.org/pdf/1807.02771.pdf
2. https://arxiv.org/pdf/1811.07661.pdf
3. https://arxiv.org/pdf/1711.06067.pdf
4. https://arxiv.org/pdf/q-bio/0605037.pdf
5. https://www.biorxiv.org/content/10.1101/713016v1
6. https://www.biorxiv.org/content/10.1101/719260v1
7. https://www.biorxiv.org/content/10.1101/467894v1
8. https://www.biorxiv.org/content/10.1101/366500v1
9. https://www.biorxiv.org/content/10.1101/466391v3
10. https://www.biorxiv.org/content/10.1101/320481v1
11. https://www.nature.com/articles/s42003-020-0874-8
12. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5311925
13. https://www.biorxiv.org/content/10.1101/768929v1.full.pdf
14. https://www.biorxiv.org/content/10.1101/751792v1.full.pdf
15. https://www.biorxiv.org/content/10.1101/2020.03.03.974766v1
16. https://www.biorxiv.org/content/10.1101/2020.05.11.089300v1
17. https://www.biorxiv.org/content/10.1101/2020.01.15.907899v1